

Stone Age babies in cyberspace

George McMurdo

Queen Margaret College, Edinburgh, Scotland, UK

Received 25 October 1995

Contents

Introduction	4
Artificial intelligence	5
Thinking about thinking	5
Modern-media mind-brain thought-experiments	6
The classification of thought	8
Consciousness of information	11
Perceiving oneself as a robot	13
Emulating a Turing machine	17
The Turing Test	17
A universal Turing machine	19
Turing’s anticipated counter-views	20
Reflections about Alan Turing and his ideas	23
Thinking in a Chinese Room	25
Inside John Searle’s Chinese Room	25
Responses to Searle’s argument	25
Reflections about John Searle and his ideas	29
The evolution of consciousness	30
Conclusion	32
References	35

Abstract.

The increasing use of computers and robots at the workplace and in peoples' personal lives, paralleled with their portrayal, interpretation and representation in the media and popular culture, has in recent years raised general awareness of some issues previously of interest mainly to philosophers of mind and consciousness. Various issues of artificial intelligence and its computational metaphors are also of more specialised relevance for information scientists, due to their bases in such things as: languages and vocabularies; categorisation and classification, and the logic of information acquisition, storage and retrieval. It may be the case that the more general acceptance of some of these ideas about human consciousness may be part of a shift in world view comparable to long-term historical precedents such as heliocentrism and natural evolution, both in the way they were initially seen to lessen humanity, and in the ways they were initially opposed and resisted.

Introduction

The concepts and issues of artificial intelligence (AI), and its attempt to understand the information processing and thinking processes of humans, are, in the mid-1990s, both popularly accessible via science fiction literature and movies, and of more specialist academic, cultural and philosophical relevance to information scientists. During the 1980s, AI also had a more practical relevance to information science, originating in the summer of 1981, when Japan's Ministry of International Trade and Industry invited leading European and American computer scientists to a conference in Tokyo to explore Japanese thinking on the 'fifth generation' computer. It announced that Japan was planning to create the next generation of computers by 1991 and wished it to be a collaborative project involving Western companies and academics. Whatever the intentions and motives, the West's computer community did not see it in that light, but rather as a declaration of commercial and intellectual war, and few attended the conference. The Americans were described as taking it as the greatest blow to scientific self-respect since the Russians launched their Sputnik in 1957 and stole the lead in the early space race. In Britain, the government set up the Alvey Committee, which recommended a £300m investment to meet the challenge with the UK's own fifth-generation computer projects.

One of the four main areas of the Alvey programme's 'enabling technologies' was intelligent knowledgebased systems (IKBSs), within which expert systems were a major component. Expert systems featured significantly in the information literature towards the end of the 1980s. It is also the case that publishing in this area has subsequently decreased significantly. This is probably partly due to a realisation that what was initially perceived as something conceptually different, to be programmed in special new AI languages like LISP and PROLOG, could also be perceived more prosaically as simply more sophisticated, but conventional, information systems, implemented more effectively with data management software and general-purpose languages like C. There is possibly also a disenchanting view that, as regards information retrieval theory, the IKBS research initiative 'reinvented the wheel' more than it contributed new knowledge. At any rate, however, implemented, an expert system can be thought of as a program which tries to encapsulate some area of human specialist knowledge. An expert system can therefore be developed in any area where a human expert is able to analyse his procedures into sets of such rules and assign statistical probabilities to them. This is also why they are called 'expert' systems, because they are programs which attempt to store and apply knowledge supplied by human experts. As such, the

range of applications of expert systems is virtually limitless, and many are in successful operation, going back to the early 1980s.

Artificial intelligence

However, as far as general AI is concerned, the same limitations could be said to apply to expert computer systems as are sometimes identified as a problem of human expert specialists. The process of specialisation in research has led to a situation where an expert is one who knows more and more about less and less. Likewise the computerised expert systems operate within narrow and highly specialised domains of knowledge and would have no ability to conduct a convincing general conversation about events, art, music, sport, etc, in a way that, for example, the three replicant characters in director Ridley Scott's 1982 film *Blade Runner* would be recognised as doing. In that movie, these robot artificial intelligences were advanced enough to be indistinguishable by visual inspection or face-to-face interrogation, and detection depended on machine measurement of a physiological pupillary response, although this was based on responses to a set of verbal cues.

Books on AI naturally tend to start by attempting to offer a definition. Not untypical (in that it side-steps the key issue of defining intelligence) is Minsky's 'the science of making machines do things that would require intelligence if done by men'¹, Similarly, Boden extends that, but also identifies the, perhaps counterintuitive, marginal importance of computer science with: 'Artificial intelligence is not the study of computers, but of intelligence in thought and action. Computers are its tools, because its theories are expressed as computer programs that enable machines to do things that would require intelligence if done by people'².

The newcomer to the philosophical issues of AI may find it disappointing that the development of a general sci-fi type 'thinking machine', able to hold conversations, is pursued only at the hobbyist level. The highly-funded national and corporate AI research initiatives are in fragmented, specialised areas, where economic benefits can be perceived.

Thinking about thinking

However, most ordinary people would sum up their interest in AI by the general question: 'Can a machine think?' (which admittedly begs a definition of what 'thinking' is) and perhaps by reference to some issue raised in sci-fi literature, TV or film. In an interesting collection encouraging participation in several 'mindexperiments' about the

¹ M. Minsky (ed), *Semantic Information Processing* (MIT Press, Cambridge, MA, 1968).

² M. A. Boden, *Artificial Intelligence and Natural Man* (Harvester Press, Hassocks, 1977).

nature of mind and consciousness, *The Mind's I*, Hofstadter and Dennett suggest a starting point for considering which of the statements in Fig. 1 seems truer³.

Initially, the first statement may seem preferable. We talk of smart people as being brains, but, rather than meaning it literally, we mean they *have* good brains. However, what if your brain were to be transplanted into a new body? Wouldn't one think of that as switching bodies? Were the vast technical obstacles to be resolved, would there be any point in a rich person with a terminal brain disease paying for a brain transplant?

Investigations of this type can be dated back to the French-born philosopher Rene Descartes (1596–1650) who aimed to found his philosophical constructions of what could be knowable by first determining what could not be doubted, since thinkers from the time of Plato had identified the unreliability of the senses. He started from a point-zero of doubting everything, and worked forward to his famous *Cogito, ergo sum* (I am thinking, therefore I exist). Since, in the process of doubting, he was thinking, then he had to exist as a thinking being. Admittedly, from this admirable cautiousness, Descartes next jumped to the conclusion that an imperfect being such as himself could not have originated such an idea, and thus it was self-evident that God exists. However, his initial discovery, that we can be certain of none of our sense-perceptions, and that we should therefore hold all our beliefs with at least some element of doubt, was confirmed in the present century by Bertrand Russell (though he qualified the observation below with the working assumption that we should not, in fact, reject a belief except on the ground of some other belief):

In one sense it must be admitted that we can never prove the existence of things other than ourselves and our experiences. No logical absurdity results from the hypothesis that the world consists of myself and my thoughts and feelings and sensations, and that everything else is mere fancy⁴.

Modern-media mind-brain thought-experiments

Frederik Schodt, author of *Inside the Robot Kingdom*⁵, about Japan's relationship with technology, has noted⁶ that the Japanese are culturally predisposed to new technology. They not only have the majority of the world's industrial robots, but find technology and related fantasy elements fun. Robot toys, images, cartoons and animations are also highly popular. By contrast, he sees the West as hampered by a certain intellectual baggage about technology, which he describes as having the 'Frankenstein

³ D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

⁴ B. Russell, *The Problems of Philosophy* (OUP, Oxford, 1959).

⁵ F. L. Schodt, *Inside the Robot Kingdom: Japan, Mechatronics, and the Coming Robotopia* (Kodansha International, Tokyo, 1988).

⁶ Channel 4 Television, *Equinox — Robotopia* (Channel 4 Television, London, 1989).

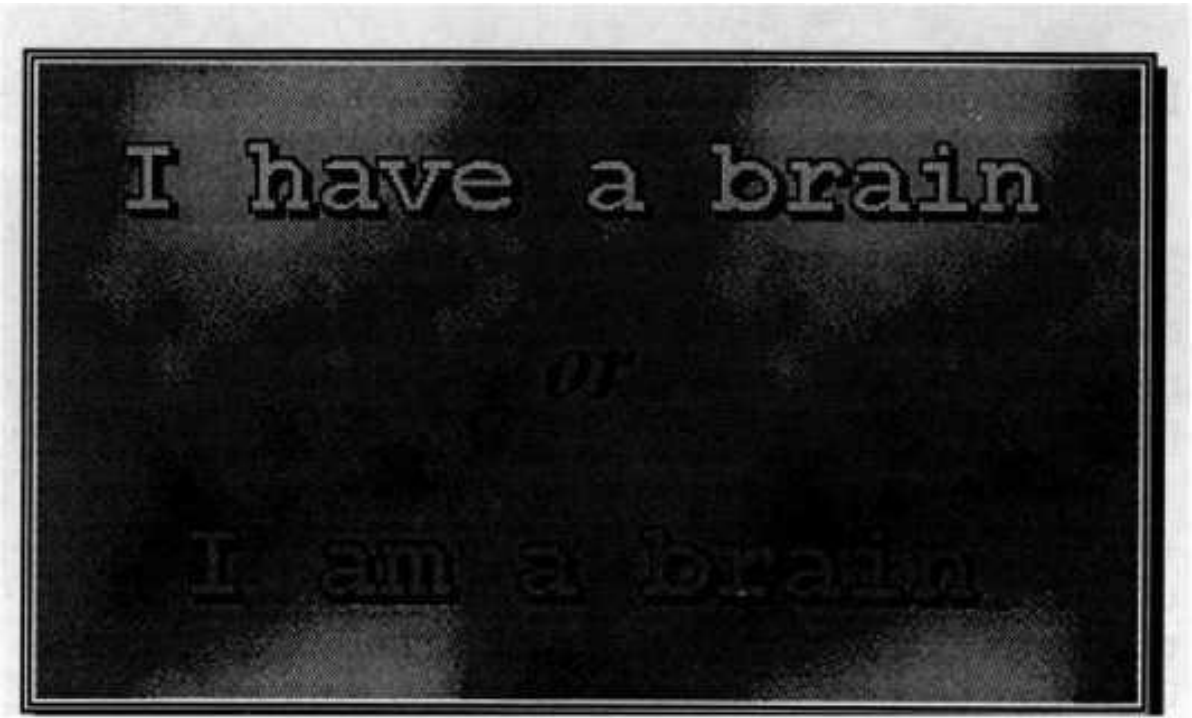


Fig. 1. Hofstadter and Dennett's introductory dilemma.

overlay', a reference to the horror movies of which the 1931 *Frankenstein*, starring Boris Karloff as the monster, is perhaps the most famous. At the present time, computer science can provide almost nothing to inform the discourse about the human mind and the possibilities and issues of artificial consciousness, largely because so very little is in fact yet known about what produces natural consciousness. At present, the most fruitful methods of exploring these issues are by imagination, literature and art, and it is noteworthy how often they indeed feature in contemporary entertainment media.

The classification of thought

Writers on the theory and practice of classification often draw attention to its philosophical bases before proceeding to the practical tasks of assigning classmarks to documents. Language itself is often the starting point, though there are many theories about the origin of speech, with agreement that the first words probably developed from primitive vocalisations. Some onomatopoeic words perhaps reflect this. Other vocalisations may have received group acceptance to become symbols for the things they represented due to the prestige of the initiator. Even today, people may adopt a saying or gesture because the person who initiates it has prestige or popularity, and words and sayings move in and out of popular usage for similar reasons.

Over time, many expressions and sayings become virtually arbitrary codes for the ideas we wish to communicate, detached from their original literal meanings. The German physicist and satirist George Cristoph Lichtenberg (1764–99) observed that ‘most of our expressions are metaphorical — the philosophy of our forefathers lies hidden within them’. Words are tools, symbols to give names to objects and ideas. Some words represent or describe, while others define relationships. There is a popular philosophical argument that it is actually impossible to think without using words. This probably is not true, since, for example, some thinking can be done with imagery. However, there is clearly some truth in this observation by Friedrich Nietzsche (1844–1900) that sometimes a name is required to allow general perception of something new:

What is originality? To see something that has no name as yet and hence cannot be mentioned although it stares us all in the face. The way men are, it takes a name to make something visible for them¹.

A child initially uses some words without knowing their meaning. Meaning is eventually acquired by discovering how these words are related to other words. Classification is one of the most significant processes in learning, as the grouping together of things which share a certain property so that some idea can be formed of the category of which the grouped things are examples. Classification and categorisation reflect the important human ability to generalise and make abstractions. Abstraction is the observation of similarity between otherwise different things. A young child acquiring the concept of a MOTOR VEHICLE makes observations that, regardless of how much motor vehicles may differ, they still have something in common. His first experience

¹ F. Nietzsche, *The Gay Science* (Vintage Books, New York, 1974).

with a motor vehicle may be with a private CAR, and he associates this with the words ‘motor vehicle’. Later, he may hear the same words associated with VANS or TAXIS, both quite different in appearance. After a series of such experiences, the child may encounter, say, a LORRY, which may never have been called a ‘motor vehicle’ in his presence, but may nevertheless designate it as such, having observed what it has in common with other motor vehicles. The child has combined aspects of past experience with his present observation of the lorry and concluded that, since it is like other motor vehicles in certain respects, the lorry is an object belonging to the category ‘motor vehicles’. Deriving a principle from past experiences is generalisation.

To a great extent, classification is something we do automatically. If we did not classify or categorise our experiences, then we would be faced with the task of interpreting each new object or event as it occurred. When we classify, however, we identify new objects and events by placing them in preconceived categories and make predictive assumptions that the new objects will behave or have an effect similar to previously experienced objects or events belonging to the same categories. Thus, as Jevons observes, our thinking processes can be seen as the results of the mind actively selecting from and organising its own experiences:

All thought, all reasoning, so far as it deals with general names or general notions, may be said to consist in classification².

Another link between information and thought, familiar to information scientists, is via the nineteenth-century mathematician George Boole, who helped to establish modern symbolic logic and from whose name derives the information retrieval term ‘Boolean operators’. One of his publications on logic was titled *The Laws of Thought*³.

However, there is also a more subtle side to this process of classification and categorisation, which at first sounds mystical and irrational. This is the sense whereby ‘reality’ is not something that is objectively ‘out there’, but rather something that is subjective, and personal. It is the sense that what we see and are able to perceive depends on what we are looking for, upon our interests or previous knowledge and experience. In a sense, we can only see what we already know. In other words, knowledge and perception of an entity is not merely an observation of what it is in isolation, but rather an awareness of its relationship with other entities. Thus the perception of the same entity may vary considerably from one individual to another, and there is no objective external reality to be seen as identical by different observers, but rather each observer brings his own reality to the construction of a perceived event. If the child above goes on to have specialist interests in motor vehicles, it may be that he will ‘see’ particular models of motor vehicles, or be able to identify some by the sounds of their engines, or even their starter-motors. The psychologist George Kelly recognised this

² W. S. Jevons, *Elementary Lessons in Logic, Deductive and Inductive* (Macmillan, London, 1870).

³ G. Boole, *An Investigation of the Laws of Thought* (Walton and Maberly, London, 1854).

effect and formulated his Personal Construct Theory to account for previously unexplained individual differences in observations by subjects in experiments. His theory asserts that reality does not directly reveal itself to us, but rather it is subject to as many alternative ways of construing it as we ourselves can invent:

Whatever nature may be, or howsoever the quest for truth will turn out in the end, the events we face today are subject to as great a variety of constructions as our wits will enable us to construe. This is not to say that at some infinite point in time human vision will not behold reality out to the utmost reaches of existence. But it does remind us that all our present perceptions are open to question and the most obvious occurrences of everyday life might appear utterly transformed if we were inventive enough to construe them differently⁴,

⁴ G. A. Kelly, *The Psychology of Personal Constructs* (Norton, New York, 1955).

Consciousness of information

It is recognised that in the ‘Nature versus Nurture’ analysis of behaviour, the human species is highly influenced by the nurturing environment and the information we assimilate from it. Likewise, our knowledge of other cultures shows us that we humans are ‘programmable’ to potentially be entirely different from what we may presently be. If we were to have grown up in a radically different culture, we would nevertheless have seamlessly learned its values. As Kelly implies, it may be difficult, but desirable in a globalised culture, to accept that one’s most deeply held beliefs are, in fact, substantially arbitrary. Indeed, not only is any of us in principle capable of taking on the mindsets of any one other example of our species, we are even programmable enough to adopt the lifestyles of other species. The fiction of *Tarzan of the Apes* stretches the imagination less than some of the factual examples in Fig. 2, drawn from Maison’s *Wolf Children*¹.

Moreover, in whatever culture we live, we are individually much determined by information, which is not only the information we have about ourselves in our minds, but also the information that others have about us, or that we believe they have about us. Our self-consciousness in social situations is thus a complex matrix of information relationships. In *The Dyer’s Hand*, W. H. Auden wrote that:

The image of myself which I try to create in my own mind in order that I may love myself is very different from the image which I try to create in the minds of others in order that they may love me².

¹ L. Maison, *Wolf Children* (NLB, London, 1972).

² W. H. Auden, *The Dyer’s Hand* (Faber, London, 1963).

WILD CHILDREN	Date of Discovery	Age at Discovery
Wolf-child of Hesse	1344	7
Wolf-child of Wetteravia	1344	12
Bear-child of Lithunama	1661	12
Sheep-child of Ireland	1672	16
Calf-child of Bamberg	cl 680	9
Bear-child of Lithuania	1694	10
Bear-child of Lithuania	9	12
Bear-girl of Fraumark	1767	18
Sow-girl of Salzburg	9	22
Pig-boy of Holland	9	?
Wolf-child of Holland	9	9
Wolf-child of Sekandra	1872	6
Wolf-child of Kronstadt	9	23
Snow-hen of Justedal	9	12
Leopard-child of India	1920	?
Wolf-child of Maiwana	1927	9
Wolf-child of Jhansi	1933	?
Leopard-child of Dihungi	9	8
Gazelle-child of Syria	1946	9
Gazelle-child of Mauritania	1960	9
Ape-child of Teheran	1961	14

Fig. 2. Cases of children reared in the wild.

The connection between these four quotations lies in the realisation that, in terms of information, an individual has both a subjective and an objective existence. In some ways this is self-evident, but in other ways may seem obscure or even absurd. The subjective existence is more real to the individual, represents fully and accurately what the individual feels himself or herself to be as a person, yet is inaccessible to the outside world. The objective existence consists of the information which has been accurately or inaccurately, consciously or unconsciously, communicated to other individuals, each of whom may construct a different interpretation.

Perceiving oneself as a robot

The archetypal film robot is surely the one in Fig. 4, created by the scientist Rotwang in Fritz Lang's 1926 *Metropolis*. The design of Lang's art deco female robot clearly also influenced the appearance of the fussy male humanoid robot See Threepio (C-3PO) in George Lucas' 1977 *Star Wars*.

Cherry¹, writing about the social effects of television as a one-way communication medium, attributes a related duality to Sartre:

... it was Jean-Paul Sartre who distinguished between these two classes of communication, in terms of 'watching' and 'being watched'. Thus, when you are in conversation with somebody you are fully aware that he or she is of the same nature as yourself, ie. human. This recognition instils into you certain attitudes and feelings which do not exist when you are quite alone ...

Sartre himself said² that people could only truly be themselves when alone, and identified three informational states of existence:

... the body exists in three modes or dimensions. I live my body; my body is known and used by another; in so far as I am an object for another, he is a subject for me and I exist for myself as known by another as a body.

Economist and author Eugene Schumacher went one better and identified four kinds of information about personal existence³, as set out in Fig. 3.

-
- (1) I — inner
 - (2) the world (you) inner
 - (3) I — outer
 - (4) the world (you) outer

These are the four fields of knowledge, each of which is of great interest and importance to every one of use. The four questions that lead to these fields of knowledge may be put like this:

¹ C. Cherry, *World Communication: Threat or Promise* (Wiley, London, 1971).

² J-P. Sartre, *L'Être et le Néant* (Gallimard, Paris, 1943).

³ E. F. Schumacher, *A Guide for the Perplexed* (Cape, London, 1977).

- (1) What is really going on in my inner world?
- (2) What is going on in the inner world of other beings?
- (3) What do I look like in the eyes of other beings?
- (4) What do I actually observe in the world around me?

To simplify in an extreme manner we might say:

- (1) = What do I feel like?
- (2) = What do you feel like?
- (3) = What do I look like?
- (4) = What do you look like?

Fig. 3. Schumacher's four fields of knowledge.

As will be seen below, the Turing Test fundamentally aims to exclude all physical and interpersonal cues from the judgement of machine intelligence. However, external appearance is clearly a strong subjective influence, and is also one which can have unexpected effects. The jerky movements of robotic street-dancers making zizzing sounds like stepper motors make them seem dehumanised and eerie. Conversely, state-of-the-art humanoid and animal animatronics is increasingly convincing where used in movies, historical tableaux, and in some marketing activities. Professor Masahiro Mori of the Robotics department of the Tokyo Institute of Technology has expressed a view of the perception of robots by humans which he calls the 'Uncanny Valley'⁴. His thesis is that the more closely the robot resembles the human being, the more affection or feeling of familiarity it can engender. However, counterintuitively, for certain ranges of operating parameters, imitation of human exteriors can lead to unexpected reversals in perception. For example, at the 1970 Osaka World's Fair a robot incorporated 29 facial muscles so that a range of expressions could be programmed. According to its designer, a smile is a sequence of facial distortions performed at a certain speed. If the speed was halved, however, a charming smile would become a frightening and uncanny grin, and the robot face drops into the uncanny valley.

Something similar has been shown to happen in perceptions of handicapped people, where the face-to-face verbalised ideas of a speech-handicapped person are demonstrated to seem more intelligent when presented in writing, or read by an actor. An interesting related exception to this has emerged through the prominence of the physically disabled physicist Stephen Hawking, who uses a speech synthesis device. As people working in the field of speech disability know, the same device is widely used. Hawking is perceived as highly intelligent, and so this positive perception is conferred on others who have the same voice, who might otherwise be perceived as less intelligent.

Joseph Engleberger, inventor of the industrial robot, has estimated that, of the 200,000 industrial robots in the world, 130,000 are in Japan⁵. Industrial robots are,

⁴ J. Reichardt, *Robots: Fact, Fiction and Prediction* (Thames and Hudson, London, 1978).

⁵ Channel 4 Television, *Equinox — Robotopia* (Channel 4 Television, London, 1989).



Fig. 4. Lang's art deco robot for *Metropolis*.

of course, non-humanoid in appearance, mainly being the ‘robot arms’ used in industrial assembly. Schodt notes⁶ that Japan is perhaps the only country where serious research is conducted towards the longterm goal of humanoid, bi-pedal robots. Professor Ichiro Kato, of Waseda University, has developed the concept of ‘my robot’, a personal assistant that will help around the home, and perhaps ‘interact with your senile grandmother’. He sees this as a 20 to 30-year goal, and suggests that such robots will functionally need to be similar to humans, in order to operate in an environment designed for humans.

Schodt suggests⁷ that, while some Japanese thinkers are becoming concerned about ‘social autism’ in their teenagers, where they relate more and more to machines, and less and less to one another, the way Japanese culture embraces robot technology and is prepared to coexist with machines and high technology, may be generally positive, compared with some Western perceptions. Treating machines more like people may lead to treating people less like machines. He notes that Japan is the only advanced industrial nation without origins in the Judaeo-Christian tradition, but is rather a pantheistic culture. For example, in the Shinto religion, inanimate objects can be sacred. Consider also the following ideas (source uncertain):

I call to mind that Buddha said ‘In this very body, six feet or so in length, with all its senses and sense impressions, and all its thoughts and ideas, in this body is the world. It is the origin of life and likewise the way that leads to the ceasing of the world... So with every action be yourself, inside yourself, listening to yourself and experiencing yourself and the richness and subtlety of what you are and what you are doing. Relax and open your mind to the subtle changes and movements deep at the roots of your being. Relax and open your mind to the awareness of the forces of the cosmos surrounding you, and flowing through you. These eternal forces are serving you all the time, and they are offering you something you have never seen before, for they were there before you were born and they will still be there after you have ceased to be. If you hear this inner music your way to discovery will be easy.

The Buddha is commonly noted, along with Socrates and Christ, for not committing his ideas to writing. Whatever the actual source, this extract has that ‘New Age’ mystical feel to it that may at first be associated with Eastern thinking, but which clicks into a different, more practical, sharpened reality if considered in the light of this quote from Masahiro Mori:

To learn the Buddhist way is to perceive oneself as a robot⁸,

⁶ Channel 4 Television, *Equinox* — *Robotopia* (Channel 4 Television, London, 1989).

⁷ Channel 4 Television, *Equinox* — *Robotopia* (Channel 4 Television, London, 1989).

⁸ Channel 4 Television, *Equinox* — *Robotopia* (Channel 4 Television, London, 1989).

Emulating a Turing machine

Alan Turing, the enigmatic pioneering British World War II cryptanalyst and computer scientist, proposed that the question, ‘Can machines think?’ be replaced by another question that can be stated without ambiguity. Turing proposed a psychological ‘imitation game’ involving three participants — a computer, a human experimenter, and a human interrogator as subject, sitting in another room, unaware of whether it is the computer or the human experimenter he is interrogating. If, after a number of trials of questioning and answering, the interrogator is unable to determine which is the human and which the computer, then the computer must be deemed to be a thinking machine.

The Turing Test

On 16 December 1994, California State University, San Marcos, in San Diego, hosted the Fourth Annual Loebner Prize Competition in Artificial Intelligence (<http://coyote.csusm.edu/cwis/loebner.contest.html>), supervised by the Cambridge Center for Behavioral Studies in Massachusetts. Ten judges ‘conversed’, using computer screens, with ten participants, five of whom were human ‘confederates’, and five of which were computer programs¹. Over a three-hour period, the judges rotated, and formed judgements about the ‘humanness’ of the ten participants as they conversed on ten restricted topics. In the end, none of the five programs was smart enough to fool any of the judges that it was human. The one that came closest was on the topic of sex, by Dr Thomas Whalen, of the Communications Research Center of the Canadian government (<http://debra.dgbt.doc.ca>), where he is employed to develop a system that will give sex advice to shy people. His program is written in C, and uses a database with only about 380 possible responses. It was never intended to appear human, and he entered the contest on a whim. Nevertheless, he came away with a cheque for \$2,000 and a bronze medal, as in Fig. 5. His winning program can be conversed with at <telnet://debra.dgbt.doc.ca:3000>.

Another participant was Dr Michael Mauldin, of Carnegie-Mellon University, who is regarded as the main developer of the Lycos system for indexing the World Wide Web (<http://lycos.cs.cmu.edu>), and who entered a program that converses on the topic of ‘Cats versus Dogs’. His program, Julia, is described on a WWW page (<http://>

¹ C. Platt, What’s it mean to be human anyway? *Wired UK* 1(1) (1995) 80–85.



Fig. 5. Loebner medal depicting A. M. Turing.

fuzine.mt.cs.cmu.edu/mlm/julia.html) which has a telnet link to telnet://julia@fuzine.mt.cs.cmu.edu, where users can log in as 'julia' and converse.

In 1991, Hugh Loebner first offered a \$100,000 prize for a program that could fool ten judges during three hours of unrestricted conversation. As this was well beyond current capabilities, he also set up the \$2,000 prize for the entry that was judged most nearly human, and allowed the programmers to designate one restricted topic for conversation. When a computer passes an unrestricted test, however, the \$100,000 prize will be awarded, and the contest discontinued.

This, in essence, is the Turing Test proposed by Alan Mathieson Turing in his 1950 paper *Computing Machinery and Intelligence*² though he then advanced it as a thought-experiment. From what Turing said in this paper, he is generally reported as predicting that by the year 2000 there would be thinking computers, though he actually expressed this somewhat differently, and said other significant things which are sometimes missed.

If Hugh Loebner has to part with his \$100,000 before 2000, it will be to a computer program which is a 'bag of tricks', as Daniel Dennett³ calls some existing programs which may mimic the human, but for spurious reasons. The most famous example of this was Joseph Weizenbaum's ELIZA (after Eliza Dolittle in Shaw's *Pygmalion*)⁴. He wrote this program to simulate a nondirective Rogerian psychotherapist conversing with a patient — a clever context, since the program required a minimal database of its own knowledge, and produced its output substantially by reflecting back the 'patient's' own dialogue. However, Dennett sees a less parodic investigation of thinking machines in a combined 'top-down' and 'bottom-up' exploration, where neuroscientists work from the bottom up, from the biological, atomic level. Cognitive scientists work from the top down, decomposing the whole person into subsystems, and component subsystems, until they finally reach components that can also be identified by the neuroanatomists.

A universal Turing machine

One of Turing's important points, which tends to get less attention than the notion of his Test, is that of a 'universal machine'. In an earlier, more mathematical, paper he demonstrated that any computing machine with primitive abilities to write and read symbols, and to do so based on instructions defined by those symbols can, given sufficient capacity, mimic any other computing machine. Thus, in a sense, all digital computers are equivalent, and *universal machines*; and any general-purpose digital computer can take on the guise of any other general-purpose digital computer.

² A. M. Turing, *Computing machinery and intelligence*. In: D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

³ J. Miller, *States of Mind* (BBC Publications, London, 1983).

⁴ J. Weizenbaum, *Computer Power and Human Reason* (Penguin, Harmondsworth, 1984).

In computing, the terminology for this is ‘emulation’, and it is this capability of embedding one architecture in another which gives rise to ‘virtual’ computers. Emulation — as opposed to simulation — means to simulate, but *exactly* in *every* respect. By contrast, simulations are always approximate. Thus computer models which simulate weather systems, or the propagation of a virus, do not get you wet, or make you ill. They are only simulating certain aspects of the phenomenon they are modelling.

The distinction that a universal machine can emulate, rather than just simulate, makes some of the arguments in what follows redundant, in that acceptance of this ability of universal machines produces a tautologous equivalence in some comparisons that are argued as different (though the term ‘emulation’ does not, in fact, arise in these arguments).

Turing’s anticipated counter-views

Turing included nine anticipated objections to his proposals (as listed in Fig. 6) and also his counterarguments to them. As will be seen, many of these overlap and are related to each other. The *Theological Objection* is that:

Thinking is a function of man’s immortal soul. God has given an immortal soul to every man and woman, but not to any other animal or to machines. Hence no animal or machine can think⁵.

Turing responds that he would find this argument more convincing if animals were classed with men, as he sees greater difference between the typical inanimate and the animate, than between man and the other animals. As a cross-religious aside, he wonders how Christians regard the Moslem view that women have no souls. His atheistic tongue seems to be somewhat in his cheek when he suggests that the quoted objection implies a serious restriction of the omnipotence of the Almighty. He asks, should we not believe that He has freedom to confer a soul on an elephant if He sees fit? Also, might we not expect that He would only exercise this power in conjunction with a mutation which provided the elephant with an appropriately improved brain to minister to the needs of this soul? The ‘*Heads in the Sand*’ *Objection* is that:

The consequences of machines thinking would be too dreadful. Let us hope and believe that they cannot do so⁶.

Theological Objection

⁵ A. M. Turing, Computing machinery and intelligence. In: D. R. Hofstadter and D. C. Dennett, *The Mind’s I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

⁶ A. M. Turing, Computing machinery and intelligence. In: D. R. Hofstadter and D. C. Dennett, *The Mind’s I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

“Heads in the Sand” Objection
Mathematical Objection
Argument from Consciousness
Argument from Various Disabilities
Lady Lovelace’s Objection
Argument from Continuity in the Nervous System
Argument from Informality of Behaviour
Argument from Extrasensory Perception

Fig. 6. Turing’s anticipated objections.

Turing suggests that while this argument is seldom expressed so openly as quoted, it affects those who think about the issue, in a way that is connected with the *Theological Objection*, and a wish to believe that Man is in some way superior to the rest of creation. He suggests that this view may be quite strong in intellectual people, since they value the power of thinking more highly than others, and are more inclined to base their belief in the superiority of Man on this power. Turing’s conclusion is that this argument is insufficiently substantial to require refutation, and that consolation may be more appropriate. The *Mathematical Objection* is based on a number of mathematical results, the best known of which is Kurt Godel’s Incompleteness Theorem that in powerful logical systems statements can be formulated which can neither be proved nor disproved, and thus there would be certain things that a computing machine embodying such principles would be unable to do. Turing does not dismiss this objection lightly, and addresses it at length, but his short answer is that although it is thus established that there are limitations to the powers of any particular machine, it has only been stated, without any sort of proof, that no such limitations apply to human intellect. For the *Argument from Consciousness*, Turing borrows a quote from a Professor Jefferson’s Lister Oration for 1949, that:

Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain — that is, not only write it but know that it had written it. No mechanism could feel (and not merely artificially signal, an easy contrivance) pleasure at its success, grief when its valves fuse, be charmed by sex, be angry or depressed when it cannot get what it wants⁷.

The *Argument from Various Disabilities* is largely the corollary to Hofstadter’s Tesler’s Theorem that AI is whatever has not been achieved yet. Turing says that

⁷ A. M. Turing, Computing machinery and intelligence. In: D. R. Hofstadter and D. C. Dennett, *The Mind’s I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

these arguments take the form: ‘I grant that you can make machines do all the things you have mentioned, but you will never be able to make one do X’. He offers a selection of features for X:

Be kind, resourceful, beautiful, friendly ... have initiative, have a sense of humour, tell right from wrong, make mistakes ... fall in love, enjoy strawberries and cream ... make someone fall in love with it, learn from experience ... use words properly, be the subject of its own thought ... have as much diversity of behaviour as a man, do something really new⁸.

Turing addresses this objection at length, dealing with many of the quoted features individually. However, his answer is generally that this objection is based on inductive views of the nature and properties of existing machines (e.g. ugly, designed for limited purposes), and associated with their very limited storage capacities. He also notes that such criticisms are often disguised forms of the *Argument from Consciousness*, and that if one does in fact present a method whereby a machine could do one of these things, one will not make much of an impression, since the method must inevitably be mechanical. He offers comparison with the parenthesis in Jefferson’s quote above. *Lady Lovelace’s Objection* is based on an extract from a memoir about her collaborator Charles Babbage’s Analytical Engine in which she states: ‘The Analytical Engine has no pretensions to *originate* anything. It can do *whatever we know how to order it to perform*’ (her italics). Turing suggests that it may not have occurred to Lovelace and Babbage that their Analytical Engine was in principle a universal machine, and that her general suggestion that machines can never do anything new, surprising or unintended is contradicted by his own experiences of working with the computers of his time. He sees this latter view that machines cannot give rise to surprises as due to a fallacy by which philosophers and mathematicians are particularly afflicted, that as soon as a fact is presented to a mind, all possible consequences of that fact spring into the mind simultaneously along with it.

The *Argument from Continuity in the Nervous System* is that the nervous system is not a discrete state machine, like a digital computer, but rather is a continuous machine, in which a small variation in the size of a nervous impulse impinging on a neuron may make a large difference to the size of the outgoing impulse. There is therefore an argument that one cannot expect to be able to mimic the behaviour of the nervous system with a discrete state system. It is slightly strange that Turing should raise this objection, since he could argue that a digital universal machine would be capable of emulating a continuous state machine. However, he bases his counter-argument on his own experience of using actual digital and analog computers, which may produce different answers for certain types of computers, but suggests that in the context of the imitation game, the approximate accuracy of answers would make it

⁸ A. M. Turing, Computing machinery and intelligence. In: D. R. Hofstadter and D. C. Dennett, *The Mind’s I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

very difficult for the interrogator to distinguish between the two. The *Argument from Informality of Behaviour* is that:

If each man had a definite set of rules of conduct by which he regulated his life he would be no better than a machine. But there are no such rules, so men cannot be machines⁹.

Turing's response to this is to suggest substitution of 'Taws of behaviour which regulate his life' for 'rules of conduct by which he regulates his life'. He believes that we can less easily convince ourselves of absence of complete laws of behaviour, than of complete rules of conduct. The method for finding such laws is scientific observation, and it is improbable that there would ever be circumstances for saying: 'We have searched enough. There are no such laws.'

The *Argument from Extrasensory Perception* is based on Turing's belief that there was overwhelming statistical evidence for telepathy. He envisages a playing of the imitation game where the human participant is as good as a telepathic receiver, and where the interrogator is asking questions such as: 'What suit does the card in my right hand belong to?'. Whereas the computer would guess at random, and get close to one in four correct, the human might score significantly higher and thus be identified. Hofstadter notes that this last objection is sometimes omitted from reprints of Turing's paper, and though he and Dennett include it, they are puzzled by Turing's motivation, if his comments are to be taken at face value and not as some sort of discreet joke. Turing's conclusion on this was that if telepathy is admitted, it would be necessary to tighten the test, and to put the competitors into a 'telepathy-proof room' to satisfy all the requirements.

Reflections about Alan Turing and his ideas

Turing's prediction about the state of thinking machines by the year 2000 was in fact in two quite distinct parts. Firstly, as better known, he believed that it would be possible to program computers to play the imitation game so well that an average interrogator would have no more than a 70% chance of making the right identification after five minutes of questioning. However, the second half of his prediction is about *the use of words* and, in some ways, is the more interesting:

The original question, 'Can machines think?' I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting

⁹ A. M. Turing, Computing machinery and intelligence. In: D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

to be contradicted. I believe further that no useful purpose is served by concealing these beliefs¹⁰.

In the present day, it would be an oversight to consider Turing's ideas in isolation from the fact that he was a homosexual, a fact that he took no particular pains to hide, especially as he grew older. His portrayal of the test as an 'imitation game', and the sense of the computer needing to second-guess the interrogator, suggests a heightened awareness of public self-presentation, and also a certain wry humour. Obviously there was no gay politics in his time, but his biographer reports¹¹ that Turing became increasingly courageous and vocal about his own sexual nature, often ignoring the advice of friends to be more cautious. Through a 1952 incident when his house was burgled, Turing indirectly, but apparently avoidably, revealed the crime of his homosexuality, and was sentenced to hormone injections which he described to a friend as causing him to grow breasts. His suicide in 1954 is sometimes attributed to harassment by British military intelligence, who may have seen him as a security risk, though it is also hypothesised that Turing may by then have seen himself as being spent as a mathematician.

¹⁰ A. M. Turing, Computing machinery and intelligence. In: D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

¹¹ A. Hodges, *Alan Turing: The Enigma* (Burnett Books, London, 1983).

Thinking in a Chinese Room

Professor of philosophy John Searle, of the University of California at Berkeley, distinguishes between what he calls ‘strong’ AI and ‘weak’ or ‘cautious’ AI. He sees weak AI as a useful modelling tool in the study of the mind, in that it enables the formulation and testing of hypotheses in a more rigorous and precise fashion. However, he rejects strong AI, in which he sees that the computer is not merely a tool in the study of the mind, but that the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have cognitive states.

Inside John Searle’s Chinese Room

Against claims that appropriately programmed computers literally have cognitive states and consciousness and that the programs thereby explain human cognition, Searle devised the *Chinese Room* thought-experiment, published in his paper *Minds, Brains and Programs*¹. We are to suppose that he is locked in a room and given a large batch of Chinese writing. Although Chinese language and symbols are meaningless to him, he is also given a second batch of Chinese script, together with a set of rules (in English, which he understands) for correlating the second batch with the first batch. By following the rules for manipulating the Chinese symbols, it can come to appear that written questions in the Chinese language are passed into the room, understood, and responses passed out again, which would be indistinguishable from those of a native Chinese speaker. Searle’s point is that this arrangement would pass a Turing Test, but that inside the Chinese Room, behaving like a computer, he actually understands nothing of what is going on.

Responses to Searle’s argument

Searle’s manuscript was made available to the main US schools of AI so that, when his article was published, no fewer than 27 responses appeared along with it. Fig. 7 lists the six Hofstadter and Dennett selected to accompany their reprinting of his article.

¹ J. Searle, Minds, brains, and programs. In: D. R. Hofstadter and D. C. Dennett, *The Mind’s I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

The replies and Searle's responses to them do not really make very enlightening reading. The *Systems*, *Brain Simulator* and *Other Minds* replies make key points, but the others given here, less so. Searle is an eloquent and stubborn debater. The AI community of the 1970s and early 1980s has been described as arrogant about the powers of its equipment. It is evident that each side has taken its position, and is at an *impasse*, mainly on the philosophically unknowable issue of artificial intentionality, and will go round and round it, much as we are told medieval philosophers would debate avidly about the numbers of angels that could dance on the head of a pin. The *Systems Reply* is:

While it is true that the individual person who is locked in the room does not understand the story, the fact is that he is merely part of a whole system, and the system does understand the story. The person has a large ledger in front of him in which are written the rules, he has lots of scratch paper and pencils for doing calculations, he has 'data banks' of sets of Chinese symbols. Now, understanding is not being ascribed to the mere individual; rather it is being ascribed to the whole system of which he is a part².

Systems Reply (Berkeley)

Robot Reply (Yale)

Brain Simulator Reply (Berkeley and MIT)

Combination Reply (Berkeley and Stanford)

Other Minds Reply (Yale)

Many Mansions Reply (Berkeley)

Fig. 7. Replies to Searle's Chinese Room argument.

Searle responds to this reply at length, but his gist is largely to reiterate that even if the individual in the Chinese Room internalises all the elements of the system, and memorises all the rules and all the Chinese symbols, he still understands nothing of Chinese, and therefore neither does the system. The *Robot Reply* is:

Suppose we wrote a different kind of program from Schank's program. Suppose we put a computer inside a robot, and this computer would not just take in formal symbols as input and give out formal symbols as output, but rather would actually operate the robot in such a way that the robot does

² J. Searle, Minds, brains, and programs. In: D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

something very much like perceiving, walking, moving about, hammering nails, eating, drinking — anything you like. The robot would, for example, have a television camera attached to it that enabled it to see, it would have arms and legs that enabled it to ‘act’, and all of this would be controlled by its computer ‘brain.’ Such a robot would, unlike Schank’s computer, have genuine understanding and other mental states³.

Searle’s answer to the *Robot Reply* is that the addition of such perceptual and motor capacities adds nothing by the way of understanding or intentionality. The *Brain Simulator Reply* is:

Suppose we design a program that doesn’t represent information that we have about the world, such as the information in Schank’s scripts, but simulates the actual sequence of neuron firings at the synapses of the brain of a native Chinese speaker when he understands stories in Chinese and gives answers to them. The machine takes in Chinese stories and questions about them as input, it simulates the formal structure of actual Chinese brains in processing these stories, and it gives out Chinese answers as outputs. We can even imagine that the machine operates, not with a single serial program, but with a whole set of programs operating in parallel, in the manner that actual human brains presumably operate when they process natural language. Now surely in such a case we would have to say that the machine understood the stories; and if we refuse to say that, wouldn’t we also have to deny that native Chinese speakers understood the stories? At the level of synapses, what would or could be different about the program of the computer and the program of the Chinese brain?⁴.

Searle insists that this reply is simulating the wrong things about the brain, that it will not have simulated what ‘matters’ about the brain, namely its causal properties, and its ability to produce intentional states. The *Combination Reply* is:

While each of the previous three replies might not be completely convincing by itself as a refutation of the Chinese room counterexample, if you take all three together they are collectively much more convincing and even decisive. Imagine a robot with a brain-shaped computer lodged in its cranial cavity, imagine the computer programmed with all the synapses of a human brain, imagine the whole behaviour of the robot is indistinguishable from human behavior, and now think of the whole as a unified system and not just as

³ J. Searle, Minds, brains, and programs. In: D. R. Hofstadter and D. C. Dennett, *The Mind’s I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

⁴ J. Searle, Minds, brains, and programs. In: D. R. Hofstadter and D. C. Dennett, *The Mind’s I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

a computer with inputs and outputs. Surely in such a case we would have to ascribe intentionality to the system⁵.

Searle says he agrees it would be rational to accept that the robot had intentionality, *as long as we know nothing more about it*. He would attribute intentionality to such a robot, pending some reason not to, but, as soon as it is known that the behaviour is the result of a formal program, he would abandon the assumption of intentionality. The *Other Minds Reply* is:

How do you know that other people understand Chinese or anything else? Only by their behaviour. Now the computer can pass behavioural tests as well as they can (in principle), so if you are going to attribute cognition to other people you must in principle attribute it to computers⁶,

This reply reverses the Turing Test and essentially asks how we know other humans have cognitive states other than assuming from behaviour. Searle is dismissive of this reply, restating his position that it could not be just computational processes and their output because the computational processes and their output can exist without the cognitive state. (A logic error in this is that while some computational processes do not give rise to cognitive states, it does not follow that no computational process can give rise to a cognitive state.) The *Many Mansions Reply* is:

Your whole argument presupposes that AI is only about analog and digital computers. But that just happens to be the present state of technology. Whatever these causal processes are that you say are essential for intentionality (assuming you are right), eventually we will be able to build devices that have these causal processes, and that will be artificial intelligence. So your arguments are in no way directed at the ability of artificial intelligence to produce and explain cognition⁷, .. ,

Searle responds that this reply redefines AI as whatever artificially produces and explains cognition, and thus trivialises the original claim that mental processes are computational processes. As such, he does not disagree with it.

⁵ J. Searle, Minds, brains, and programs. In: D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

⁶ J. Searle, Minds, brains, and programs. In: D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

⁷ J. Searle, Minds, brains, and programs. In: D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

Reflections about John Searle and his ideas

The main weakness in Searle's *Chinese Room* argument is the one exposed by the *Systems Reply*. Searle induces his reader to identify with the person in the room, shuffling pieces of paper and understanding no Chinese, but this is not how brains and minds work. What we have consciousness of is not our mental processes themselves, but of the results of those processes. It has been said that we have 'underprivileged access' to the goings-on in our own minds. His final comments, after dealing with the replies, can perhaps best be summed up by these comments:

But the equation 'mind is to brain as program is to hardware' breaks down at several points, among them the following ... the program is purely formal, but the intentional states are not in that way formal. They are defined in terms of their content, not their form. The belief that it is raining, for example, is not defined as a certain formal shape, but as a certain mental content with conditions of satisfaction, a direction of fit, and the like. Indeed the belief as such hasn't got a formal shape in this systematic sense, since one and the same belief can be given an indefinite number of different syntactic expressions in different linguistic systems... as I mentioned before, mental states and events are literally a product of the operation of the brain, but the program is not in that way a product of the computer⁸.

If he is indeed arguing against the equation of mind- to-brain as program-to-hardware, then he has been at cross-purposes, since this is an incorrect analogy. When representing any particular machine, computer programs are not distinguishable from computer hardware in this way. When a reprogrammable computer is running a particular program, it becomes that particular machine. The program is simply a convenient method of configuring the machine. Likewise, non- reprogrammable — say mechanical — machines can nevertheless have algorithms which map their behaviour. Hofstadter and Dennett surmise that:

A thinking machine is as repugnant to John Searle as non- Euclidean geometry was to its unwitting discoverer, Gerolamo Sacheri, who thoroughly disowned his own creation. The time — the late 1700s — was not quite ripe for people to accept the conceptual expansion caused by alternate geometries. About fifty years later, however, non-Euclidean geometry was rediscovered and slowly accepted⁹,

⁸ J. Searle, Minds, brains, and programs. In: D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

⁹ D. R. Hofstadter and D. C. Dennett, *The Mind's I: Fantasies and Reflections on Self and Soul* (Penguin, Harmondsworth, 1982).

The evolution of consciousness

Of course, the issue of where human intentionality itself originates from remains a real one, with the replacement of supernatural explanations by scientific ones. The current model is the evolutionary one, and the psychologist Nicholas Humphrey has presented an interesting thesis¹² that consciousness — perhaps peculiar to, or particularly developed in, human beings — is the result of evolutionary adaptation in a species which depends on co-operation and being able to second-guess each other's thoughts. The neo-Darwinian biologist Richard Dawkins, known for applying computational principles to the study of biology, offered his views on the possibility of machine consciousness:

Could you ever build a machine that was conscious? I know that I am conscious. I know that I am a machine. Therefore it seems to me and I know that there's nothing special I mean perhaps I could say that I have faith, but I think I almost know that there's nothing in my brain that couldn't in principle be simulated in a computer. So if you took the extreme policy of building a computer that was an exact simulation of a human brain, doing everything that a human brain does, a point-for-point mapping from human brain anatomy to computer hardware, then of course such a machine would have to be conscious. It would be conscious in just the same sense as that I know that I am conscious. That's a different matter from saying that it will ever be done. It would be formidably difficult to do that precise reconstruction of a computer version of a human brain³,

He proceeded to consider if it could happen that computers in actual commercial use might become sufficiently complicated that consciousness would emerge. In the evolutionary framework, he presumed the answer would be no, unless it became commercially important for computers to second-guess what humans or other computers were thinking. Dawkins made another TV appearance on BBC 2's *Late Show* during the 1993 UK Science Week, to present his *Viral Theory of Human Behaviour*. In this, he notes the contemporary metaphor of computer viruses, and that both computers

¹ N. Humphrey, *The Inner Eye* (Faber and Channel 4 Books, London, 1986).

² Channel 4 Television, *The Inner Eye — The Ghost in the Machine* (Channel 4 Television, London, 1986).

³ Channel 4 Television, *The Inner Eye — The Ghost in the Machine* (Channel 4 Television, London, 1986).

and biological cells provide environments suitable for infection by parasitic code. He suggests that there may also now be media viruses, where misinformation takes on a life of its own, and replicates and spreads. He wonders what other program-friendly environments there may be ripe for parasitic exploitation. Human brains, he decides, seem a good candidate, especially young children's brains, while they most need to learn.

Conclusion

Sigmund Freud said that both Darwin's theory of evolution and his own psychoanalytic theories had resulted in an affront to mankind's naive egoism. In *The Structure of Scientific Revolutions*¹, Thomas Kuhn showed that science advances against stubborn resistance by those who have a stake in the *status quo*, noting that the view of the heliocentrism advanced by Polish astronomer Nicolaus Copernicus (1473–1543) made few converts for almost a century after his death. (Indeed in a very recognisable way, it is still resisted. We all still think of the sun as rising and setting, rather than of the horizons rotating, since that is a more useful day-to-day working model, even though it is false.) Max Planck once said that a new scientific truth does not triumph by convincing its opponents but by outliving them and being accepted by a new generation. It may be that an evolving popular perception of information will in a significant way end the Copernican revolution in world-view and be a millennial event in the sense of spanning the 500-year period from Gutenberg's printing press to the date predicted by Alan Turing for discourse about machine intelligence.

Robert Wright's *Time* article² *The evolution of despair* has the subtitle *A new field of science examines the mismatch between our genetic makeup and the modern world, looking for the source of our pervasive sense of discontent*. Compare the sense of this subtitle with McLuhan's concluding proposal in his introduction to *Understanding Media*:

Examination of the origin and development of the individual extensions of man should be preceded by a look at some of the general aspects of the media, or extensions of man, beginning with the never-explained numbness that each extension brings about in the individual and society³.

Other writers, such as Marx, also clearly have a claim on the idea that aspects of modern society produce alienation. R. D. Laing has made the point that a modern 20th-century child is nevertheless still biologically born as a 'Stone Age baby' but that with the benefits of the Western educational system 'by the time the new human being is fifteen or so, we are left with a being like ourselves. A half-crazed creature, more or less adjusted to a mad world'⁴. Wright's piece begins with this quote:

¹ T. S. Kuhn, *The Structure of Scientific Revolutions* (University of Chicago Press, Chicago, 1970).

² R. Wright, *The evolution of despair*, *Time (International)* 146(9) (1995) 36–12.

³ M. McLuhan, *Understanding Media: The Extensions of Man* (RKP, London, 1962).

⁴ R. D. Laing, *The Politics of Experience and the Bird of Paradise* (Penguin, Harmondsworth, 1967).

I attribute the social and psychological problems of modern society to the fact that society requires people to live under conditions radically different from those under which the human race evolved ... — THE UN-ABOMBER⁵,

The Unabomber is a US self-styled anarchist, presumed to be a white male living in Northern California, who since 1978 has planted or mailed sixteen package bombs that have killed three and wounded 23, most of his victims being in universities or airlines, whence the ‘un’ and ‘a’ initial three characters of his name (<http://www.fbi.gov/unabomb.htm>). The Unabomber has a 35,000-word manifesto titled *Industrial Society and the Future* (<http://vip.hotwired.com/special/unabom>) copies of which, in April 1995, he mailed to the *New York Times*, the *Washington Post*, and *Penthouse* magazine, apparently (by the FBI’s chronology) at the same time as to his last bomb victim to date. He said that in exchange for publication of his essay, he would no longer use his bombs to kill people. On 19 September 1995, the *Washington Post* published the entire manifesto as a pull-out section jointly funded with the *New York Times*. Their publishers issued a statement that they had been advised by Attorney General Janet Reno, and the FBI, who have a \$1,000,000 reward for information leading to identification, arrest and conviction of the person(s) involved.

Wright’s article explores the work of a growing group of scholars — evolutionary psychologists — who are anticipating the coming of a new field of ‘mismatch theory’, which would study maladies resulting from contrasts between the modern environment and the ‘ancestral environment’. Wright suggests that while we may not share the Unabomber’s approach to airing a grievance, the grievance itself may feel familiar. However, when the Unabomber complains that the ‘technophiles are taking us all on an utterly reckless ride into the unknown’, he is probably correct in the latter part of the assertion, but probably mistaken (in the typical conspiracy-theory paranoid way) in the former belief that any individuals have any significant control over the fate of our species in its ride into the unknown.

Richard Dawkin’s *Late Show* piece developed to demonstrate epidemiological aspects of religions, which may spread ‘faith viruses’ whose symptoms might be deep inner certainty of truth, without supporting evidence. He suggested that various ‘fads’ for dress, behaviour, popular sayings, etc, may be cultural viruses, and show the same epidemiological spread as computer and DNA viruses. This is how he concluded, however, perhaps suggesting that in addition to being a neo-Darwinist, he is also becoming a neo-McLuhanist:

In today’s culture of globally-linked computer and media networks what new media-viruses might be emerging? Perhaps culture as a whole is a collection of such viruses and its evolution is getting faster and faster as the means of transmission become electronic and interconnected. Whatever

⁵ R. Wright, The evolution of despair, *Time (International)* 146(9) (1995) 36–12.

the future holds, we can be sure of one thing. The more programfriendly our environment becomes, the more we will have to get used to living with viruses of all sorts and until we have developed a way of talking about and identifying these new viruses, there will be no hope of vaccination.

References

The Ted K Archive

George McMurdo
Queen Margaret College, Edinburgh, Scotland, UK
Stone Age babies in cyberspace
Received 25 October 1995

Journal of Information Science, 22 (1) 1996, pp. 63–77.

doi.org/10.1177/016555159602200107

Correspondence to- G. McMurdo, Department of Communication and Information Studies, Queen Margaret College, Clerwood Terrace, Edinburgh EH12 8TS, Scotland, UK. Tel: +44 1313 173505. Fax: +44 1313 164165. E-mail: cimmu@cis.qmced.ac.uk.

www.thetedkarchive.com